

NASA's Earth Science Data Systems Standards Process Experiences

Richard E. Ullman
NASA/ Goddard Space Flight Center
Greenbelt, MD USA
richard.e.ullman@nasa.gov

Yonsook Enloe
SGT Inc.
Greenbelt, MD USA
yonsook@mindspring.com

Abstract—NASA has impaneled several internal working groups to provide recommendations to NASA management on ways to evolve and improve Earth Science Data Systems. One of these working groups is the Standards Process Group (SPG). The SPG is drawn from NASA-funded Earth Science Data Systems stakeholders, and it directs a process of community review and evaluation of proposed NASA standards. The working group's goal is to promote interoperability and interuse of NASA Earth Science data through broader use of standards that have proven implementation and operational benefit to NASA Earth science by facilitating the NASA management endorsement of proposed standards. The SPG now has two years of experience with this approach to identification of standards.

We will discuss real examples of the different types of candidate standards that have been proposed to NASA's Standards Process Group, such as OPeNDAP's Data Access Protocol, the Hierarchical Data Format, and Open Geospatial Consortium's Web Map Server. Each of the three types of proposals requires a different sort of criteria for understanding the broad concepts of "proven implementation" and "operational benefit" in the context of NASA Earth Science data systems. We will discuss how our Standards Process has evolved with our experiences with the three candidate standards. (*Abstract*)

Keywords - component; best practices; standards; community

I. BACKGROUND

NASA's Earth Science Data Systems Working Group's (ES-DSWG) Standards Process Group (SPG) is one of several standing agency working groups charged with developing recommendations for the on-going evolution of NASA's Earth science computer data systems as a whole. The purpose of these working groups is to provide a way for data systems practitioners within the agency to provide input and feedback to help guide the agency in the adoption of computer data systems technologies, software, practices and standards. The ES-DSWG's were formed in January 2004. [1]. The SPG's particular focus is to advise NASA decision makers on community views of the best practice with respect to the use of standards [2]. We believe that use of standards will maximize the value of NASA's investment in Earth science data systems by capitalizing on the "network effects" of using common interfaces. Network effects have the benefit of lowering the cost to use the data, increasing the opportunities for data inter-use or data interoperability, and increasing the integration of NASA data in communities of practice. We use the term "communities of practice" or simply "communities" to include

various stakeholders and affected constituents. Example communities include science discipline groups, users of similar applications, data systems developers, NASA Earth science mission planners, Earth science educators, and others. Membership in each "community" often overlaps the others.

Often, adoption of standards by can take the form of top down decision-making, wherein management chooses the standards that the enterprise will use based upon trade studies or market forces. But in the domain of Earth Science Data Systems, the expertise for making wise choices in standards adoption is often at the practitioner level and this is the primary reason that NASA has empanelled the SPG. In studying the question of how to best advise NASA on the best standards to use, we have seized upon a few principles. The first is that to enable interoperability, we do not require homogeneous systems, but rather we need coordination at the interfaces. The second is that communities of practice have home grown solutions to share and interoperate with data that have evolved to meet their particular needs. The solutions may include software application interfaces, data and metadata model conventions, data and information identification, common data services, formats, and other related technologies. And the solutions that are built by a community have a strong appeal to members. They feel an ownership stake in the success of their way of doing business.

The SPG process borrows from the Internet Engineering Task Force (IETF), the concept of the RFC (Request for Comment). The success of the internet RFC process shows that the publication of practices that demonstrate benefit leads to growing use through broader community adoption and adoption across communities.

Standards that we are interested in include any components or technology practices that if documented and more widely used would promote: easier sharing or exchanging of data among distributed partners and users; distributed systems development and sharing of software and technical expertise; reducing the cost of developing or maintaining a system; increasing the use of scientific data products; increasing interoperability and enhancing innovation, collaboration, and computing performance. For each candidate standard, we assess if there is a community of use that has experience in implementation and has demonstrated operational readiness and also has the leadership necessary to promote the advantage of wider use.

II. THE SPG STANDARDS PROCESS

A. Organization

The Standards Process Group (SPG): This is the decision-making board of the standards process composed of part time permanent members from NASA's program office, Earth science mission projects, Earth science funded data systems awardees and representative from other agencies.

Technical Working Groups (TWG): These are groups commissioned by the SPG to conduct public review and evaluation of candidate standards, related implementations and operational experience. Membership in a TWG is temporary and partially drawn from the Standards Process Group membership and partly drawn from technical area experts and/or community members.

A. Path to RFC

The term "RFC" stands for "Request for Comment". The content of an RFC is either a technical note or a proposed standard. A technical note is any information that the submitter considers significant to the use of a practice within NASA's Earth science programs.

RFCs can come from any NASA stakeholder source including individuals that may be associated with or represent NASA's Earth science funded activities, industry or users of Earth science data. In some cases, the SPG may solicit an RFC. Other times, members of the community will bring forward an RFC to formalize NASA recognition or broaden use of standards that are used in their community. The requirements for an RFC will be the same in each case. We require the RFC proposer to describe the practice or specification in technical detail or else provide references that describe the standard. The proposal must identify the community of use and citations of successful implementation and evidence of operational readiness must be provided.

B. Path to Community Endorsement

Figure [1] shows the steps from an RFC to endorsement as a NASA Standard. The process is characterized by technical analysis, open public review and demonstration that the proposal "works". The first step is for the SPG to perform an initial screening and characterization. A TWG is assigned and a schedule is set, taking into consideration NASA need dates and support commitments. Also, any RFC must have two or more implementations before it can advance to draft status.

The TWG invites the community by means of email announcements to comment on the specification, operational readiness of implementations, and the usefulness of the technology and particularly to address questions formulated by the TWG. The TWG also identifies key stakeholders that are likely to have particular experience with the technology and solicits their opinion. The TWG reports to the SPG and the SPG makes recommendations for the final status of the RFC.

The role of the TWG is central to the review process. Because there is a wide variety of technologies that might contribute to interoperability or data interuse, the circumstances of each RFC are different. The TWG is the place where we

wrestle with questions such as: What are the expectations for review of a specification? What evidence will show implementation? What does readiness for operational use mean for this RFC? And, what does suitability for use mean for this technology? The answers to these dictate the custom tailoring of our process to the particular RFC. The TWG forms particular questions to guide its evaluation of the RFC and to solicit reviews from and opinion of the community of practice.

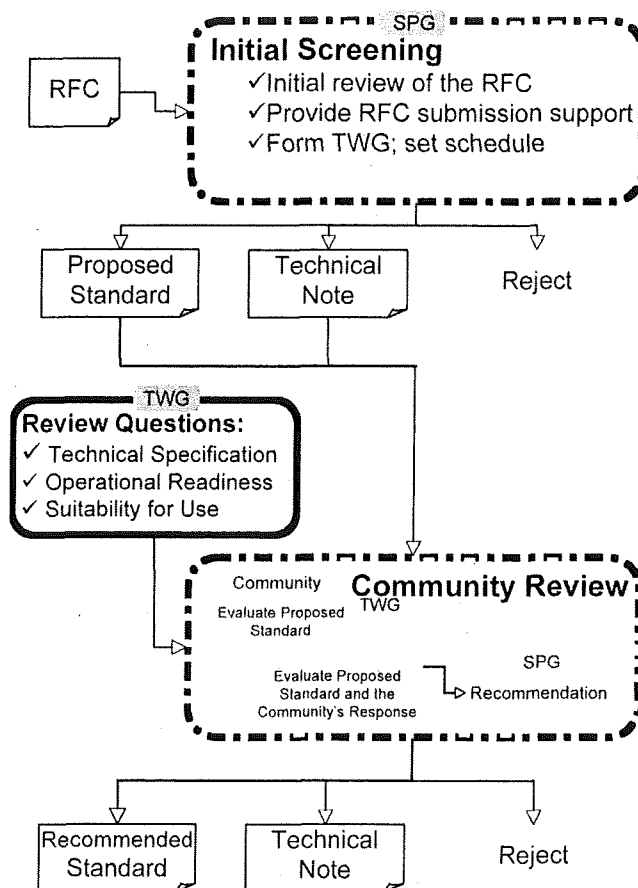


Figure 1. Standards Endorsement Process

The three types of reviews, the specification review, operational readiness review, and the suitability for use review, are conducted at the same time. For the specification review, the TWG asks reviewers to answer questions about accuracy of the specification in the RFC and to evaluate the significance of at least two implementations. The TWG must determine whether the implementations are independent and interoperable uses of the practice. For example, if the RFC proposes a format for a particular class of science product, demonstration of the use of that format by two separate "implementing organizations" would be considered two implementations. In the operational readiness evaluation, the TWG focuses on operational readiness of the implementations. Not only must the standard be demonstrated to work, but also the standard must be shown to work under conditions that are judged by the

TWG and ultimately the SPG to assure that the implementations of the standard are robust enough to be ready for operational contexts of NASA data and NASA stakeholder users. The TWG also conducts a "usefulness for purpose" review, asking users to evaluate whether the proposed technology is suitable for a named purpose.

III. STANDARDS EXPERIENCE

To date, the SPG has processed three proposed RFCs. We now describe the experience with each.

A. OPeNDAP DAP 2 RFC

In summer 2004, the OPeNDAP Group submitted the DAP (Data Access Protocol) v2.0 specification as a candidate community standard. The DAP 2 is a "home grown" standard, not adopted by any de rigueur standards organization and future "from scratch" implementations are expected. In this first run through of the Standards Process, the Standards Process worked very well. The community was very responsive with extensive and comprehensive answers to the detailed review questions about the specification, implementations, and operational experience. The reviews were painless for the SPG to obtain. Why? In retrospect, we think this is because the OPeNDAP community was very cohesive and engaged. The OPeNDAP Group and key community users provided very strong leadership for the community and were successful in "getting out the vote". The author of the RFC was very responsive in correcting specification errors found during the reviews. The result? The SPG recommended endorsement of the DAP 2 as a NASA Earth Science standard.

Judging from the increased interest in using DAP 2.0 by the mission-success element of the Goddard DAAC and MODAPS and the increased success of proposals that propose to use OPeNDAP, the SPG endorsement of DAP 2.0 has been successful in lowering the barriers to the acceptance of DAP 2. The OPeNDAP Group was asked about the benefit of the SPG process for the DAP 2.0, they noted "The OPeNDAP Board of Directors singled this activity out as one of the most important for the past year. They felt that the benefits were well worth the (low) costs" [3].

The OPeNDAP Group's positive experience with the SPG has encouraged The HDF Group to submit the HDF (Hierarchical Data Format) v5 specification as a candidate community standard to the SPG. They hope to get the NASA community endorsement of the HDF5 before submitting the HDF5 specification document to international standards organizations.

B. Hierarchical Data Format v5 RFC

The HDF5 is a homegrown standard, originally developed by the NCSA, not adopted as a standard by any other external organization. A single implementation, in the form of shared software libraries, exists. Due to the high cost of implementing this extensive specification, NASA is unlikely to fund a completely independent implementation of this standard. For the purposes of NASA's Standards Process, an independent use of the shared software libraries by independent projects is considered to be an independent implementation for these

reviews. The HDF5 RFC documented independent use of the software libraries by multiple independent projects.

During the technical specification review, it was difficult to get many reviews. Because there was only one implementation of the software libraries, not that many people had experience implementing the HDF5 from the specification. The SPG recognized that in this kind of a situation where there is a common shared software library, less numerous technical specification reviews will need to be allowed. Some users returned comments on the usability of HDF5 for their purposes. After some extensive discussion, the SPG decided that review comments on the usability or suitability of a proposed technology for a purpose does need to be part of the assessment, especially for a technology that the users are directly exposed to and use, and should be used in the overall evaluation. This RFC is currently being evaluated for operational readiness.

C. OGC Web Map Service RFC

The Open Geospatial Consortium (OGC) Web Map Service RFC was submitted to the Standards Process Group as a candidate community standard. This specification is already an OGC and an International Standards Organization (ISO) standard. Future, completely new implementations are expected. Again, it was difficult to get many technical specification reviews. However, this time, the reason was different. Reviewers wondered why we are asking for a specification review when this specification was already extensively reviewed and vetted by the OGC and ISO. After extensive discussion, the SPG decided that technical specification reviews for candidate standards already adopted by external standards organizations are not needed.

The review of operational experience was also difficult. "Operational experience" means something very specific in the NASA context. Some of the reviewers mentioned that they were serving hundreds of users each day with thousands of accesses to images, but didn't consider themselves operational. But clearly the WMS could be used in an operational environment based on the actual number of users, data volume, and hours of daily operation. The formal designation of "Operational" sometimes comes much later in the software life cycle within NASA. The SPG decided to ask for reviews of operational readiness instead of operational experience. The SPG recommended endorsement of the OGC WMS as a NASA Earth Science Recommended standard.

IV. EXPERIENCE SUMMARY

From our experience with the three RFCs described above, we found that all three types of reviews; the technical specification review, the usefulness for purpose review, and the operational readiness review, can all be done at the same time. Many times the same organizations have experience in all three areas and would prefer to send all the reviews at the same time instead of being asked sequentially. We have agreed that "Operational Readiness" is more important than "Operational Experience" or else the standards adoption becomes trailing edge. We also have found that our original description of

"community" which was based on "science discipline community" has now evolved to also include those who have experience with implementation of a particular proposed standard. Already we are experiencing increased adoption of a standard that has been endorsed by the Standards Process Group. As additional community standards are endorsed and recommended for use for NASA funded data systems, we expect the endorsed standards to have significant impact on the evolution of NASA's Earth science data systems.

REFERENCES

- [1] NASA ES-DSWG. "Earth Science Data System Working Group." Internet: <http://seeds.gsfc.nasa.gov/> Jan. 12, 2006 [Apr. 20, 2006].
- [2] NASA Standards Process Group. "NASA Earth Science Data Systems SPG - Portal." Internet: <http://spg.gsfc.nasa.gov> Jan. 06, 2006 [Apr. 20, 2006].
- [3] J. Gallagher. "Not Your Standard Deal: OPeNDAP's Positive Experience with NASA ESDS SPG." Presented at the NASA ES-DSWG SPG meeting, San Diego, June 15 2005. Internet: <http://spg.gsfc.nasa.gov/spgfolder/events/esdswg-meeting-june-15-2005/opendap.pdf> [Apr. 20, 2006].